

SEQUENCE LISTING

SEQ ID NO:1

X-Sun-Data-Type: text

5 X-Sun-Data-Description: text

X-Sun-Data-Name: text

X-Sun-Charset: us-ascii

X-Sun-Content-Lines: 11

menin.cDNA 1.28, 2772 bases

10 1 GGTGTCCGGA GCCGCGGACC TAGAGATCCC AGAAGCCACA GCGCAGCGGC
51 CCGGCCCCGCC ACTATTTCCA GGCTCTGCGG GCCAGGGGCC GCCGCCCACC
101 GCCCGCCGCC ATGGGGCTGA AGGCCGCCCA GAAGACGCTG TTCCCGCTGC
151 GCTCCATCGA CGACGTGGTG CGCCTGTTTG CTGCCGAGCT GGGCCGAGAG
201 GAGCCGGACC TGGTGCTCCT TTCCTTGGTG CTGGGCTTCG TGGAGCATTT
15 251 TCTGGCTGTC AACC GCGTCA TCCCTACCAA CGTTCCCGAG CTCACCTTCC
301 AGCCCAGCCC CGCCCCGAC CCGCCTGGCG GCCTCACCTA CTTTCCCGTG
351 GCCGACCTGT CTATCATCGC CGCCTCTAT GCCCGCTTCA CCGCCCAGAT
401 CCGAGGCGCC GTCGACCTGT CCCTCTATCC TCGAGAAGGG GGTGTCTCCA
451 GCCGTGAGCT GGTGAAGAAG GTCTCCGATG TCATATGGAA CAGCCTCAGC
20 501 CGCTCCTACT TCAAGGATCG GGCCCACATC CAGTCCCTCT TCAGCTTCAT
551 CACAGGCACC AAATTGGACA GCTCCGGTGT GGCCTTTGCT GTGGTTGGGG
601 CCTGCCAGGC CCTGGGTCTC CGGGATGTCC ACCTCGCCCT GTCTGAGGAT
651 CATGCCTGGG TAGTGTTTGG GCCCAATGGG GAGCAGACAG CTGAGGTCAC
701 CTGGCACGGC AAGGGCAACG AGGACCGCAG GGGCCAGACA GTCAATGCCC
25 751 GTGTGGCTGA GCGGAGCTGG CTGTACCTGA AAGGATCATA CATGCGCTGT
801 GACCGCAAGA TGGAGGTGGC GTTCATGGTG TGTGCCATCA ACCCTTCCAT
851 TGACCTGCAC ACCGACTCGC TGGAGCTTCT GCAGCTGCAG CAGAAGCTGC
901 TCTGGCTGCT CTATGACCTG GGACATCTGG AAAGGTACCC CATGGCCTTA
951 GGGAACCTGG CAGATCTAGA GGAGCTGGAG CCCACCCCTG GCCGGCCAGA
30 1001 CCCACTCACC CTCTACCACA AGGGCATTGC CTCAGCCAAG ACCTACTATC
1051 GGGATGAACA CATCTACCCC TACATGTACC TGGCTGGCTA CCACTGTCGC
1101 AACC GCAATG TGCGGGAAGC CCTGCAGGCC TGGGCGGACA CGGCCACTGT
1151 CATCCAGGAC TACA ACTACT GCCGGGAAGA CGAGGAGATC TACAAGGAGT
1201 TCTTTGAAGT AGCCAATGAT GTCATCCCCA ACCTGCTGAA GGAGGCAGCC
35 1251 AGCTTGCTGG AGGCGGGCGA GGAGCGGCCG GGGGAGCAAA GCCAGGGCAC
1301 CCAGAGCCAA GGTTCCGCCC TCCAGGACCC TGAGTGCTTC GCCCACCTGC
1351 TGCGATTCTA CGACGGCATC TGCAAATGGG AGGAGGGCAG TCCCACGCCCT

1401 GTGCTGCACG TGGGCTGGGC CACCTTTCTT GTGCAGTCCC TAGGCCGTTT
1451 TGAGGGACAG GTGCGGCAGA AGGTGCGCAT AGTGAGCCGA GAGGCCGAGG
1501 CGGCCGAGGC CGAGGAGCCG TGGGGCGAGG AAGCCCGGGA AGGCCGGCGG
1551 CGGGGCCCCAC GGCGGGAGTC CAAGCCAGAG GAGCCCCCGC CGCCCAAGAA
1601 GCCAGCACTG GACAAGGGCC TGGGCACCGG CCAGGGTGCA GTGTCAGGAC
1651 CCCCCCGGAA GCCTCCTGGG ACTGTGCTG GCACAGCCCG AGGCCCTGAA
1701 GGTGGCAGCA CGGCTCAGGT GCCAGCACCC GCAGCATCAC CACCGCCGGA
1751 GGGTCCAGTG CTCACCTTCC AGAGTGAGAA GATGAAGGGC ATGAAGGAGC
1801 TGCTGGTGGC CACCAAGATC AACTCGAGCG CCATCAAGCT GCAACTCACG
1851 GCACAGTCGC AAGTGCAGAT GAAGAAGCAG AAAGTGTCCA CCCCTAGTGA
1901 CTACACTCTG TCTTTCCTCA AGCGGCAGCG CAAAGGCCTC TGAACACTG
1951 GGGACTTCGG ACCGCTTGTTG GGGACCCAGG CTCCGCCTTA GTCCCCAAC
2001 TCTGAGCCCA TGTTCTGCCC CCAGCCCCAA GGGGACAGGC CTCACCTCTA
2051 CCCAAACCCT AGGTTCCCGG TCCCGAGTAC AGTCTGTATC AAACCCACGA
2101 TTTTCTCCAG CTCAGAACCC AGGGCTCTGC CCCAGTCGTT AGAATATAGG
2151 TCTCTTCTCC CAGAATCCCA GCCGGCCAAT GGAAACCTCA CGCTGGGTCC
2201 TAATTACCAG TCTTTAAAGG CCCAGCCCCT AGAAACCCAA GTCCTCCTC
2251 GGAACCGCTC ACCTAGAGCC AGACCAACGT TACTCAGGGC TCCTCCAGC
2301 TTGTAGGAGC TGAGGTTTCA CCCTTAACCC AAGGGAGCAC AGGTCCCACC
2351 TCCAGCCCGG GGAGCCTAGG ACCACTCAGC CCCTAGGAGT ATATTTCCGC
2401 ACTTCAGAAT TCCATATCTT GCGAATCCAA GTCCTCGCC CCAAATAACT
2451 TCAGTCCTGC TTCCAGAATT TGGAATCCT AGTTTCCTCT CCTTCGTATC
2501 CCGAGTCTGG GACACAAAAC TCCGCCCCCA GCCTATGAGC ATCCTGAGCC
2551 CCGCCCTCTT CTGACGAAA CTGGCCCCCG ATCAGAGCAG GACCTCCCTT
2601 CCGACCCTCT GGGAACCTCC CAGAGGTCCA GCCCATCTCG GAGCATCCCG
2651 GAGGAAATCT GCAGAGGGGT TAGGAGTGGG TGACAAGAGC CTGATCTCTT
2701 CCTGTTTTGT ACATAGATTT ATTTTTCAGT TCCAAGAAAG ATGAATACAT
2751 TTTGTTAAAA AAAAAAAAAA AA

SEQ ID NO:2

Mu.prot sequence 610 amino acids

1 MGLKAAQKTL FPLRSIDDVV RLFAAELGRE EPDLVLLSLV LGFVEHFLAV
51 NRVIPTNVPE LTFQPSAPD PPGGLTYFPV ADLSIIAALY ARFTAQIRGA
101 VDLSLYPREG GVSSREL VKK VSDVTWNSLS RSYFKDRAHI QSLFSFITGT
151 KLDSSGVAFA VVGACQALGL RDVHLALSED HAWVVFPGNG EQTA EVTWHG
201 KGNEDRRGQT VNAGVAERSW LYLKGSYMRC DRKMEVAFMV CAINPSIDLH
251 TDSLELLQLQ QKLLWLLYDL GHLEYPMAL GNLADLEELE PTPGRPDPLT
301 LYHKGIASAK TYRDEHIYP YMYLAGYHCR NRNVREALQA WADTATVIQD

351 YNYCREDEEI YKEFFEVDAND VIPNLLKEAA SLLEAGEERP GEQSQGTQSQ
401 GSALQDPECF AHLLRFYDGI CKWEEGSPTP VLHVGWATFL VQSLGRFEGQ
451 VRQKVRIVSR EAEAAEAEEP WGEEAREGRR RGPRRESKPE EPPPPKKPAL
501 DKGLGTGQGA VSGPPRKPPG TVAGTARGPE GGSTAQVPAP AASPPPEGPV
551 LTFQSEKMKG MKELLVATKI NSSAIKLQLT AQSQVQMKKQ KVESTPSDYTL
601 SFLKRQRKGL

SEQ ID NO:3

The genomic sequence for the *MEN1* gene

mu_genomic.fa 9180 bp

1 CTGGTCTTGA ACTCCTGGCC TCAAGCAATC CTCCTGCTTC AGCTTCCCAA
51 AGTGTGTGTA TTACAGGCAT GAGCCTGGCA TGAAGTTGAC ACTATTGAGA
101 TATACTGGTC AGGTATTTTG TGAATGTCC CTCAACTCTG TTTTGCCAGA
151 TGTCTTCTCA TGATTAGAGG AGAGTTATAA ATTTTGAGGA AAATCCAGAG
201 AGGTGAAGAG GTGAAGTAGG GCAGAAATTT AATCTGTTTT ATTTACTGCT
251 ATATACCGAG TGTCTGGAAC TTGGCCCATG GTAAGTACCA AAAATCTGTT
301 TTTTTTGAAT GAATAAGCAA ATAAATGAGT GACCGTGGAA ATTTAGTATT
351 ATTTCAAAGT TTCAAAGCGT TGTTGATACA GGCCAGGCAC AGTGGCTCAC
401 ACCTGTAATC CCAGCACTTT TGGAGGCCGA GGTAGGAGGA TCACTTGAGG
451 TCAGGAGTTC GAGACCAGCC TGACCAACAT GGTGACACCC CTGTCTCTAC
501 TAAGTAAAAT AAAAAAATTA GCCAAGTGTG GTGGCAGGCA CCTGTAATCC
551 CGGCTACTTG GGAAGCTGAG GCAGAAGAAT CACTTGAACC TGGGAGGCAG
601 AGGTTGCAGT GAGCCGAGAT CACCCCACTG CACTCCAGCC TGAGTGACAG
651 AGCGAGACTC TGTCTCAAAA CAAATAAACA AATAACTACT CTTTGGCCGG
701 GTAAGGTGGT TCACGCCTGT AATTTTAGCA CTTTGGGAGG CTGAGGCGGG
751 CAGATCACTT GAGGTTAGGG GTTCGAGACC AGTCTGGCCA ACATGGTGAA
801 ACCCCATCTC TACTTAAAT AAAAAAGTT TTCTGGGTGT GGTGGCGGAC
851 GCCTATAATC CCAGCTACTT GGGACTTTTT TTTAAGACGG AATCTCACTC
901 TGTGCCCCAG GCTGGAGTGC AGTGGCAAGA TTCTGGCTCA CTGAAGCCTC
951 CGCCTCCCAG GTTCAAGGGG ATCCCCGCG CCTCAGCCTC CCAAGTAGCT
1001 GGAATCCCT GTCTCTGCAA AAAAAAAAAA AAAAAAAAAAC AAAAAATATA
1051 TATATATATA TATATGTGTG TGTGTGTGTG TGTGTGTGTG TTATATGTAT
1101 ATATATTTAT GTATATGCAC ATACACACAA AATTAGGCGG GAGTGGTGGC
1151 GCACGCCTGT GATCACAGCT ACTCGGGAGG CTGAGGCACG AGAATCGCTT
1201 GAGCCCGTGA AGTCGAGGCT GCAGTGAGCC CAGATCGAGC CACTGCATTC
1251 CAGCCTGGGC GAAAGAGAAA GACCGTGTCT CAAAACAAAC AAACAAAAGC
1301 TACTCTTAGC ACGTGTTAGA GTATCTCGCG GGCGGAAGTG GGAACGAGT
1351 GCTGCACACA GAGTAGGCAT CTTTATATGT TAACAGACAC TGATACCCAG

1401 CTAAAGCGGC TGAACACATT TACTCTCTGG CAGTGTTTAA AAGTATCTGT
1451 TTTTCTCATA TTGTTTTATT TTAATTTTTT CTGGATCAAG CAACCTGATC
1501 TTTTTCCTCA TAACTTGCCG ACCGACCCGT GACAGCAAAA CCGGCAGAAG
1551 CTCGGCGACC TCCCACCCCG AGTCTGCAGG TAGTGCCCCC GGA CTACTACATT
1601 TTCCAGAAGG CACTGCGGGC ACGCTTCCTG CCTGGTCCGGC CTGAAGGGAA
1651 GGGCCAATCC CTGAGTATCT CGGGAAGGAG GTGTCCGGAG CCGCGGACCT
1701 AGAGATCCCA GAAGCCACAG CGCAGCGGCC CGGCCCCGCA CTATTTCCAG
1751 GCTCTGCGGG GCAGGGGTGG GCCCAGACTC CACTTCCCGG CGGGTAGTGC
1801 GACCCTAGGG GCGGGACTTC ATGTCCCAGC AGGCTCCGGG CGGCGTGCGC
1851 CGCGGTGCCT AGTGTGGGAT GTAAGCGCGG AGGTGGGCGA GGGGGACCGA
1901 GGCCAGGACT CTCCTTGGGG TTTGGGGGCT TGACCTGGGT GCGCTTTCTG
1951 GACAGACTTT ACAGCCCCCG GGGGCACAGT CGTAGAGAGG GGGCGGGGCG
2001 GCCATTGGGG CTCCTCATTG GGGTGCTTGG GCGCACCCC ATCGGGTACC
2051 GGGCGTCCCG GAATTGTGGG GGACAAAAAG GCTCTGCAGT CTCGGCTGAG
2101 GGGTCTCACC GACAAAAGAG GGGAAGCCGG TGAGCAGAGG CTGAAGAGGG
2151 TGGGGAAGCA GGGGAGCTGT GCGTGTGTCG GGGCGGGTGG AACCTTAGCG
2201 GACCCTGGGA GGAGGCTCCC CGGCCGAACC TGCCCGACCC TCCCTCCCCC
2251 GGCTTGCTT GCAGGCCGCC GCCCACC GCCCGCCATG GGGCTGAAGG
2301 CCGCCCAGAA GACGCTGTTT CCGCTGCGCT CCATCGACGA CGTGCTGCGC
2351 CTGTTTGCTG CCGAGCTGGG CCGAGAGGAG CCGGACCTGG TGCTCCTTTC
2401 CTTGGTGCTG GGCTTCGTGG AGCATTTTCT GGCTGTCAAC CGCGTCATCC
2451 CTACCAACGT TCCCGAGCTC ACCTTCCAGC CCAGCCCCGC CCCCAGCCCC
2501 CCTGGCGGCC TCACCTACTT TCCCGTGGCC GACCTGTCTA TCATCGCCGC
2551 CCTCTATGCC CGCTTCACCG CCCAGATCCG AGGCGCCGTC GACCTGTCCC
2601 TCTATCTCG AGAAGGGGGT GTCTCCAGCC GTGAGCTGGT GAAGAAGGTC
2651 TCCGATGTCA TATGGAACAG CTCAGCCGC TCCTACTTCA AGGATCGGGC
2701 CCACATCCAG TCCCTCTTCA GTTTCATCAC AGGTTGGAGC CCAGTAGGTG
2751 GGAATCTTAT CCATGACCCA CTTCTTCAAA ACCCTCCATG GTTTACAGAA
2801 CCCTTTTAAG AACTGTAAGC CTTGTGAGGT TCGGCAGGTG TTATTTTCTT
2851 CTTTGCAGTT GGGAACTGA AGCCCAGAGA GGGGAAATGA TATGCCAAAG
2901 TCACACACGG CATGGCAGGG CTGGAAGTGA AGCCTGATCA CTTGGCTCCA
2951 AATCATCAAC CTCACCTCTG CCCCCTCAGC ACCCCCACCC TTGCCACTGA
3001 ACAGCTACAG GAGTTCTAAG CATGAGACAC AGAGGGCGGC AGCAGATTTA
3051 GGGGGCAAGA AGATGAAATT GGGCTGCATT TGAGGCAGTT AAACAAAATA
3101 ATGGCTATGA AGATTTTTTT TTTTTTTTTT TTTTGAGACA GGGTCTCACT
3151 CTGTCCCCCA GGCTGGAGTG CAGTGGTGTG ATCATGGCTC ACTGCAGCCT
3201 CAGTCTCCCT GGGCTCAGAG ATCCTCCAAC CTCAGCCTCC TGAGTAGCTG
3251 AGAGTACAGG CATGCACCGT GGTGCTGGTT AATTTTTTGT ATTTTTTTTG

3301 TAGAGATGGT GTCTCACTAT GTGGCCCAGA CTGGTCTTGA ACTCTTGGGC
3351 TCAAGTGATC TGCCCGCCTC AGTCTCCCAA ATGCTGGGAT TACAGGTGTG
3401 AGCCACCGCA ACTGGTGGCC TATGAAAATT TTTTTTTTTT TTCAGACGGC
3451 GTCTCACTCT GTCGCCCAGG CTGGAGTGCA GTGGTGCAAT CTCGGCTCAC
5 3501 TGCAAGCTCT GCCTCCTGCT TTCATGCCAT TCTCCTGCCT CCTGCCTCAG
3551 CCTCCTGAGT AGCTGGGACT ACAGGAGCCT GCCACCATGC CTGGCTAATT
3601 TTTTTTTGGA TTTTATAGTAG AGACGAGGTT TCACCATGTT AGCCAGGATG
3651 GTCTCGATCT CCTGACCTCG TGATCCGCCC GCCTTGGCCT CCCAAAGTGC
3701 TGGGATTACA GCGTGAGCC ACCGCACCTG GTCAAAAATG TTTGAGACAG
10 3751 AGAAGGGGCT TGACCTCAAA AGGCTTAAGA GTCAGGGCTT GCAAAGAGCT
3801 TTGCACCAAG CCCGGTTGAC TGGCAATCCC ATCCTGGTGT GCCATATTGA
3851 GAAGGAATCA GAGGCTGCTT CTCAGCTTAG CAGGAAAAGA GTGCAGAGAT
3901 AAATGAGGGT TATTTGTTGG TGGGTGTATA GCCAGAGAGT GTTGGCCAGC
3951 GTCCTGTTTT TGCCATTCTT GTTTTAACCT AGTAAGTGCA GTAAAATGGA
15 4001 ATCCCTAAAT CCATAGAATA TATAATAGAG TTGCAGAGAA AGACGAGGTA
4051 GGGCCAAAGG CTGGGTCAGC TACAGGATAT CCAGAAAGGT ATCTTGTTGG
4101 ACATAGAGGG TGTAACAGG GAGAGAGTCT TTGAACACGT GGGAGGGAAG
4151 GGATGGAGGG ATAGTGGGCA GGAGAATCTG AGGTTGGGTC ACAGGCTTGG
4201 AAAGGGAGTG GGAGGGAGTG TGGCCCATCA CTACCTGGCC CCTTTCCCCA
20 4251 TGTTAAAGCA CAGAGGACCC TCTTTCATTA CCTCCCCCTT CCACAGGCAC
4301 CAAATTGGAC AGCTCCGGTG TGGCCTTTGC TGTGGTTGGG GCCTGCCAGG
4351 CCCTGGGTCT CCGGGATGTC CACCTCGCCC TGTCTGAGGA TCATGCCTGG
4401 GTAGTGTTTG GGCCCAATGG GGAGCAGACA GCTGAGGTCA CCTGGCACGG
4451 CAAGGGCAAC GAGGACCGCA GGGGCCAGAC AGTCAATGCC GGTGTGGCTG
25 4501 AGCGGGTATT GTTCCCTCCC CCCAGCCTTG TCCCCTTCAT ACTGTAGTAG
4551 CCCAAGCCAC CCAAGGGACT CCATTTTCTT GGGCCACACC CCTTTCTTCC
4601 CATCACCACC CACATAGGAA GGAAGACAG AAGAGCCCCT TTTCTTGGCT
4651 GTCATTCCCT GAAGCAGGCA CAGGGTGGG CATCATGAGA CATAATGATC
4701 TCATCCCCC CTAAGAGCTG GCTGTACCTG AAAGGATCAT ACATGCGCTG
30 4751 TGACCGCAAG ATGGAGGTGG CGTTCATGGT GTGTGCCATC AACCTTCCA
4801 TTGACCTGCA CACCGACTCG CTGGAGCTTC TGCAGCTGCA GCAGGTGAGG
4851 GCTGAGCCAA TGGGGCAGGA CTGGGCTAGG CCAGACTTGA CTTGCTGTGG
4901 GACCCTGGGC AGGGGCACTT TCCCTTCCTG AGCTTCAGCT TCCCCTCCTG
4951 GAAAAATGGG TTAGTAATTC CTGGCCTGGC CTTTCCCAGG GCTCTTGGGA
35 5001 GAGTAGAATT GAGATGTGAA ATTGCTTTGA CTCCATTAAA GGGCTGGTCC
5051 CAGAATTTTG GCCCTTCCAC ATGGTGGGTG GTCCCTGTTG GTTCTGACCC
5101 CCACCTCTGC CCGATAGGCT AAGGACCCGT TCTCCTCCCT GTTCCGTGGC
5151 TCATAACTCT CTCCTTCGGC TCCTAGAAGC TGCTCTGGCT GCTCTATGAC

5201 CTGGGACATC TGGAAAGGTC AGTAGAGGGA AGTGGCCAGG CTGCGCCTGG
5251 TGAGGCCGGG GGGCTGGGTG GCAGCCTGAA TTATGATCCT TTCCTAGGTA
5301 CCCCATGGCC TTAGGGAACC TGGCAGATCT AGAGGAGCTG GAGCCCACCC
5351 CTGGCCGGCC AGACCCACTC ACCCTCTACC ACAAGGTGGG GGCATCTAAG
5401 GAGGGTGCAG AAGGGAGACC CTAACAGTGG CTGAGGCAGG GGCCCTCATC
5451 TGGGCAGATG AGAAGAGAAC TTTGTGTGTT GGGGGGTATC GCCCATCCAG
5501 TCTCACTTTG TGTCAACTGT GTGCAGAATC AGTTCAGTCA GGGCTGTCTG
5551 AGGGGTGTCC AGGGTTCCCC AGCCTGGGAG TGGCAGGGGC TGCATTTGTC
5601 CCCTCAGCCC TGCCTTTTCT GCCACTGCTT ACTGTCCTTC CTGGAGTATA
5651 ACAGAGGTCA AATGTGGTAG GAGCACTGAA GAGGGGGTGT TCACTTGGTG
5701 GGTGTAGGTG GGGAGGAGGG CCATTGGGCT GGGCTTGAAA GTCTTTGGTG
5751 ATGTGTAGAA GAGTGTCTGA GAAAGAGAAG GGCCCTGAGC TCGGAGGGCA
5801 GGCCCCACCC CTGCAGTCTG CCCCAGGCCT CAGCCAGCAG TCCTGTAGAC
5851 CCAGGGAGGA GACCAGGTAG AAGGGCTGGC AGCGAGTGGA GGTGGGAGTG
5901 GAGATGGAGA GGACTCCCTG GGATCTTCCT GTGGCCCCCTT CTGGGTGTGC
5951 CCTGGTGGGG CATTGTGCC AGCAGGGCAG CTGGGGCTGC CTCCCTGAGG
6001 ATCCTCTGCC TCACCTCCAT CCAGGGCATT GCCTCAGCCA AGACCTACTA
6051 TCGGGATGAA CACATCTACC CCTACATGTA CCTGGCTGGC TACCACTGTC
6101 GCAACCGCAA TGTGCGGGAA GCCCTGCAGG CCTGGGCGGA CACGGCCACT
6151 GTCATCCAGG AGTGAGGATC CCCCTACTAG GGCCTGCAGC CTGTCCTTTC
6201 TTCCCCTCCA TCAGTTTCCA ACCACCCTCG TCCAGGACTG AGGCCTGGCT
6251 CCCACGCCCC ATCCCCTTTC CATCCAGTCC CTAGGCAGCA AGGCCACCAT
6301 TACCCAGGAG GTAGGGACCC TGATTAAGGT GTCACATCTT TCCCTCCCTC
6351 CCCTCTCCTC CTAATTTTTT TTTTCTCAGA ACAGTCTCAA ATCTCCAATG
6401 TTTAACCACC ATCATCCAGC AGTGGGACTT CCACCCTCGG CCCCATGCCC
6451 CCCTCCTCAT TCTTGCTTTC TTCCTCTGGG CTGACCCAGA CAGCATCATT
6501 TTGCAGTGAG GACCCACCT ACTCCCCAG CCCCTGGGGG CTCCATCCCC
6551 CGCCAGGTCC CTGGGGCTAC CCCCAGTGGT GAGACCCCTT CAGACCCTAC
6601 AGAGACCCCA CTGCTCTCAC AGCTACAACT ACTGCCGGGA AGACGAGGAG
6651 ATCTACAAGG AGTTCTTTGA AGTAGCCAAT GATGTCATCC CCAACCTGCT
6701 GAAGGAGGCA GCCAGCTTGC TGGAGGCGGG CGAGGAGCGG CCGGGGGAGC
6751 AAAGCCAGGT GAAAGGCTGG AGCTCCAGCC TGTGTCCAGC CTCCCACCTG
6801 GACAGGGCTC CTTTCCACAG GGCCATGGGG GCTGCATGTA CGGGATTAGG
6851 GATGGCAGGA GGAAGGTGGC CCTGAGCAGA CAGCTATGTT CCCTTTTGCT
6901 ATAAGTGAAG TCCTGGGCCC ACGTTGGACG GGACTGAAGG TATTTTAGAG
6951 GTTTCTACCC TGTGCCTTCA GTTTCATGGC CAGACTCCCT CCCTCAGCTG
7001 AGGGGTGGAG GTAGGGATGG TACGTCCTGG CTATGGATTG GCTTTATAAA
7051 AGGAAAGAGG TTCTAAGAAT GTTCCCAACC TATGCTTACC TTTTCTGGAG

7101 CCAGGGGTCT TTGCCTAGGT GGGGGGCCTG GCCTGTGCCC TCTGCTAAGG
7151 GGTGAGTAAG AGACTGATCT GTGCCCTCCC TTCCCCCTCG TCCAGGGCAC
7201 CCAGAGCCAA GGTTCGCCCC TCCAGGACCC TGAGTGCTTC GCCCACCTGC
7251 TGCGATTCTA CGACGGCATC TGCAAATGGG AGGAGGGCAG TCCCACGCCT
5 7301 GTGCTGCACG TGGGCTGGGC CACCTTTCTT GTGCAGTCCC TAGGCCGTTT
7351 TGAGGGACAG GTGAGGGACA GCTGCACAGA GGTCTGGGCA CTACAGGTGG
7401 TGACAGCAGC CACGGGCTTG TCAGACTTTT CTGGCCCAGG GGCAGCATCT
7451 GCCCATCCCC TTCGGTGCCG ATGGGACTGA GACCCCCTGG GTGGGATGGG
7501 ATGGCCAGAG CAGGGTCCTG GAGTTCCAGC CACTGGCCGG CAACCTTGCT
10 7551 CTCACCTTGC TCTCCCCACT GGCCCAGGTG CGGCAGAAGG TGCGCATAGT
7601 GAGCCGAGAG GCCGAGGCGG CCGAGGCCGA GGAGCCGTGG GGCGAGGAAG
7651 CCCGGGAAGG CCGGCGGCGG GGCCCACGGC GGGAGTCCAA GCCAGAGGAG
7701 CCCCCGCCG CCAAGAAGCC AGCACTGGAC AAGGGCCTGG GCACCGGCCA
7751 GGGTGCAGTG TCAGGACCCC CCCGGAAGCC TCCTGGGACT GTCGCTGGCA
15 7801 CAGCCCGAGG CCCTGAAGGT GGCAGCACGG CTCAGGTGCC AGCACCCGCA
7851 GCATCACCAC CGCCGGAGGG TCCAGTGCTC ACTTTCCAGA GTGAGAAGAT
7901 GAAGGGCATG AAGGAGCTGC TGGTGGCCAC CAAGATCAAC TCGAGCGCCA
7951 TCAAGCTGCA ACTCACGGCA CAGTCGCAAG TGCAGATGAA GAAGCAGAAA
8001 GTGTCCACCC CTAGTACTA CACTCTGTCT TTCCTCAAGC GGCAGCGCAA
20 8051 AGGCCTCTGA ACTACTGGGG ACTTCGGACC GCTTGTGGGG ACCCAGGCTC
8101 CGCCTTAGTC CCCCAACTCT GAGCCCATGT TCTGCCCCCA GCCCAAAGGG
8151 GACAGGCCTC ACCTCTACCC AAACCCTAGG TTCCCGGTCC CGAGTACAGT
8201 CTGTATCAAA CCCACGATTT TCTCCAGCTC AGAACCAGG GCTCTGCCCC
8251 AGTCGTTAGA ATATAGGTCT CTTCTCCCAG AATCCCAGCC GGCCAATGGA
25 8301 AACCTCACGC TGGGTCCTAA TTACCAGTCT TTAAGGCCC AGCCCCTAGA
8351 AACCCAAGCT CCTCCTCGGA ACCGCTCACC TAGAGCCAGA CCAACGTTAC
8401 TCAGGGCTCC TCCCAGCTTG TAGGAGCTGA GGTTTCACCC TTAACCCAAG
8451 GGAGCACAGG TCCCACCTCC AGCCCGGGA GCCTAGGACC ACTCAGCCCC
8501 TAGGAGTATA TTTCCGCACT TCAGAATTCC ATATCTTGCG AATCCAAGCT
30 8551 CCCTGCCCCA AATAACTTCA GTCCTGCTTC CAGAATTTGG AAATCCTAGT
8601 TTCCTCTCCT TCGTATCCCG AGTCTGGGAC ACAAACTCC GCCCCAGCC
8651 TATGAGCATC CTGAGCCCCG CCCTCTTCCT GACGAACTG GCCCCGGATC
8701 AGAGCAGGAC CTCCTTCCG ACCCTCTGGG AACCTCCCAG AGGTCCAGCC
8751 CATCTCGGAG CATCCCGGAG GAAATCTGCA GAGGGGTTAG GAGTGGGTGA
35 8801 CAAGAGCCTG ATCTCTTCCT GTTTGTACA TAGATTATT TTTCAGTTCC
8851 AAGAAAGATG AATACATTTT GTTAAAAAA ATATAAAGCG CAAGTCCATG
8901 TTTATCTGGG AAATTGGGGA TGGGGCGGGG AGTGGAGCGC CCCTTCTTCC
8951 CTTTGTCTTC TGGCTCCCGG GACTTTGCGC TCCCTACCTG TGGAGCGCGA

9001 GCGACAGTGG CGGCGGAAGG ACGTAGGCTC CGCCCCGGCC TTGGGGCTTC
9051 CCCC GCGCCG CCGAGGGCCC GTCCCGCGGG CGCCTCCTCC CGGACTGGCG
9101 GTGGGGCATC CCNGGGCGCG GCCCGCCCC CGGGCTTCAG CCCC GCCCCC
9151 GCGGCTTCAG AGCCACGGGC GCCCGCCCCG